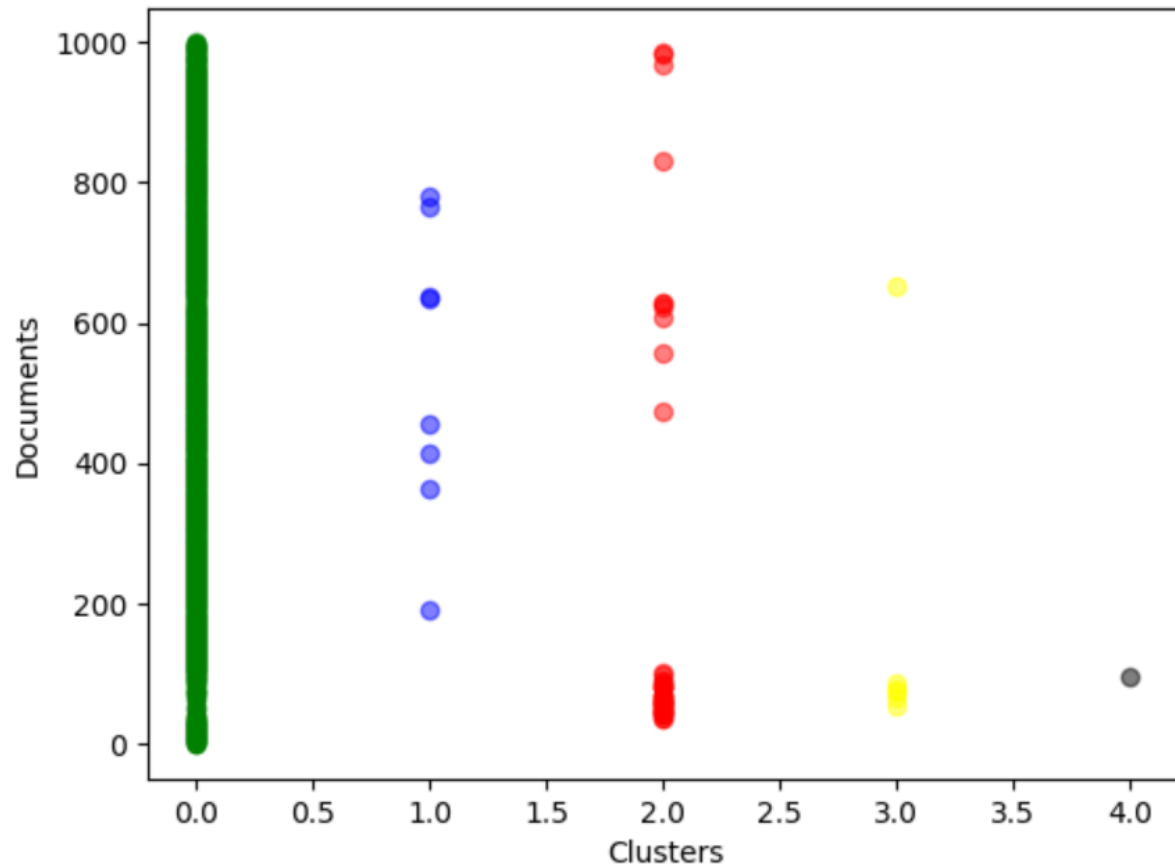**BONUS QUESTION:**

Visualization:

This visualization is done on Kmeans from Task1.

The X-axis shows the clusters. I have 5 clusters with cluster_ids = [0,1,2,3,4]

The Y-axis shows the document ids.

The plot shows which documents ids belong to which cluster.

Cluster 0 has the most number of documents, whereas cluster 4 is relatively sparse.

****
****